Measurement of the Relative Branching Fraction of $B_s^0 \to J/\psi f_0(980), f_0(980) \to \pi^+\pi^-$ to $B_s^0 \to J/\psi\phi, \phi \to K^+K^-$

B. Abbott
Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman, OK, USA

A measurement of the relative branching fraction of $B_s^0 \to J/\psi f_0(980), f_0(980) \to \pi^+\pi^-$ to $B_s^0 \to J/\psi\phi, \phi \to K^+K^-$ is presented. The decay mode $B_s^0 \to J/\psi f_0(980)$ is an interesting mode since it is a CP-odd eigenstate which could be used in CP-violating studies. Using approximately 8 fb^{-1} of data recorded with the D0 detector at the Fermilab Tevatron Collider, a relative branching fraction of $0.210 \pm 0.032\,(\text{stat}) \pm 0.036\,(\text{syst})$ is found.

I. INTRODUCTION

The CP-violating phase in B_s^0 mixing mixing, has been measured [1, 2] using $B_s^0 \to J/\psi\phi$ decays. The measured absolute value is larger than predicted by the Standard Model (SM) [3], but is statistically consistent with it. The decay products in $B_s^0 \to J/\psi f_0(980)$ are in a CP-odd eigenstate and can provide a more direct measurement of this CP-violating phase. Measuring this CP-violating phase using $B_s^0 \to J/\psi f_0(980)$ decays mode can aid in reducing its uncertainty.

Based on estimates the relative branching fraction should be large. Using hadronic D_s^+ decays, Stone and Zhang [4, 5] estimated the relative width to be:

$$R \equiv \frac{\Gamma(B_s^0 \to J/\psi f_0(980); f_0(980) \to \pi^+\pi^-)}{\Gamma(B_s^0 \to J/\psi\phi; \phi \to K^+K^-)} \approx 0.20. \quad (1)$$

The LHCb collaboration has reported [9] a first measurement of $R = 0.252^{+0.046+0.027}_{-0.032-0.033}$. The Belle collaboration has made a measurement of the branching fraction $\mathcal{B}(B_s^0 \to J/\psi f_0(980); f_0(980) \to \pi^+\pi^-) = (1.16^{+0.31}_{-0.19}\,(\text{stat.})^{+0.15}_{-0.17}\,(\text{syst.})^{+0.26}_{-0.18}(\text{N}_{\text{B}_s^{(*)}\bar{\text{B}}_s^{(*)}})) \times 10^{-4}$ [10]. The CDF collaboration has also measured the relative branching fraction and finds R=0.257 $\pm$ 0.020 (stat) $\pm$ 0.014 (syst) [11]. This article provides a new measurement of the relative branching fraction using the D0 detector collecting data at the Fermilab Tevatron Collider.

This note provides a new measurement of the relative branching fraction from D0.

II. RELATIVE BRANCHING FRACTION

To determine an absolute branching fraction, various efficiencies, branching fractions, and cross sections need to be known, as well as the integrated luminosity. However, by measuring a relative branching fraction, several terms common to both the $B_s^0 \to J/\psi f_0(980)$ branching fraction and the $B_s^0 \to J/\psi\phi$ branching fraction cancel giving:

$$R = \frac{\mathcal{B}(B_s^0 \to J/\psi f_0(980); f_0(980) \to \pi^+\pi^-)}{\mathcal{B}(B_s^0 \to J/\psi\phi; \phi \to K^+K^-)} = \frac{N_{B_s^0 \to J/\psi f_0(980)} \times \varepsilon_{reco}^{B_s^0 \to J/\psi\phi}}{N_{B_s^0 \to J/\psi\phi} \times \varepsilon_{reco}^{B_s^0 \to J/\psi f_0(980)}}. \quad (2)$$

All that is required to measure a relative branching fraction are the relative yields and the relative reconstruction efficiencies of the two decay modes, $\varepsilon_{reco}^{B_s^0 \to J/\psi\phi}$ and $\varepsilon_{reco}^{B_s^0 \to J/\psi f_0(980)}$.

III. SELECTION CUTS

A. Analysis Cuts

The data set of an integrated luminosity of approximately 8 fb^{-1} was divided into four periods corresponding to different detector configurations called RunIIa, RunIIb1, RunIIb2 and RunIIb3.

The initial sample of $B_s^0 \to J/\psi f_0(980)$ was found by first reconstructing $J/\psi \to \mu^+\mu^-$ candidates by requiring that two oppositely charged muon candidates with transverse momentum $p_T > 1.5$ GeV form a common vertex.

Since the D0 detector has a limited ability to separate kaons from pions, all reconstructed tracks not associated to a J/ψ are considered for reconstructing $f_0(980)$ and ϕ candidates. The tracks are assigned the pion mass when searching for $B_s^0 \to J/\psi f_0(980)$ and the kaon mass when searching for $B_s^0 \to J/\psi\phi$. Two tracks with a minimum p_T of 300 MeV, having an invariant mass 0.7 GeV $< M_{\pi^+\pi^-} <$ 1.2 GeV, and being consistent with coming from a common vertex were considered as $f_0(980)$ candidates. Finally, the $\mu^+\mu^-\pi^+\pi^-$ candidates were required to have a common vertex and have an invariant mass between 5.0–5.8 GeV.

Similar requirements were applied to the initial sample of $B_s^0 \to J/\psi\phi$ candidates. The only different requirements were that 0.91 GeV $< M_{K^+K^-} <$ 1.05 GeV and the $\mu^+\mu^-K^+K^-$ candidates were required to have an invariant mass between 5.0–5.8 GeV. Due to the invariant mass requirements on $M_{\pi^+\pi^-}$ and $M_{K^+K^-}$, two tracks cannot be considered both a $f_0(980)$ and a ϕ candidate. The final data sample was then formed by applying the additional requirements:

- All runs without optimal performance of muon, silicon microstrip and central fiber trackers are omitted .
- All events that only fired a trigger that required muons with a large impact parameter were removed.

J/ψ selection:

- Both muons are required to be detected as a track segment in either one or three layers of the muon system and be matched to a central track.
- At least one muon must be detected as a track segment in three layers of the muon system.
- Both muons must have at least one hit in the silicon microstrip tracker.
- 2.9 GeV $< M_{\mu^+\mu^-} <$ 3.2 GeV

$f_0(980)$ (ϕ) selection:

- Both pions (kaons) from the $f_0(980)$ (ϕ) candidate must have at least 2 hits in the central fiber tracker.
- Both pions (kaons) from the $f_0(980)$ (ϕ) candidate must have at least 2 hits in the silicon microstrip tracker.
- Both pions (kaons) from the $f_0(980)$ (ϕ) candidate must have at least 8 hits total in the silicon microstrip tracker and the central fiber tracker.
- The momentum of the leading pion (kaon) from the $f_0(980)$ (ϕ) candidate must be greater than 1.4 GeV.
- $f_0(980)$ (ϕ) candidate p_T must be greater than 1.6 GeV.

B_s^0 selection:

- 0.91 GeV $< M_{\pi^+\pi^-} <$ 1.05 GeV (when searching for $J/\psi f_0(980)$.)
- 1.01 GeV $< M_{K^+K^-} <$ 1.03 GeV (when searching for $J/\psi\phi$.)
- $p_T(B_s^0) >$ 5.0 GeV
- Proper decay length [13], L, significance, $L/\sigma(L) > 5$, where $\sigma(L)$ is the uncertainty on the proper decay length.

B. Boosted Decision Trees

It is known that boosted decision trees (BDT) [14, 15] are a powerful tool for separating signal from background. Signal and background samples are used to train the BDT and a discriminant is determined for each event. By making a selection on the value of the BDT discriminant, the signal to background ratio can be vastly improved. We use the Monte Carlo (MC) PYTHIA program [16] to generate B_s^0 and the EVTGEN program [17] to simulate its decay. Two MC background samples were produced: a prompt sample (directly produced J/ψ) and an inclusive sample (all decay processes $B_s^0 \to J/\psi + X$). A MC signal sample of $B_s^0 \to J/\psi f_0(980)$ events was then used to train the BDT on both the prompt and inclusive background. A BDT discriminant was found for both the prompt and inclusive sample and used in the analysis. A total of 36 different kinematic variables were used to train the BDT consisting of isolation variables, transverse momentum of the daughters

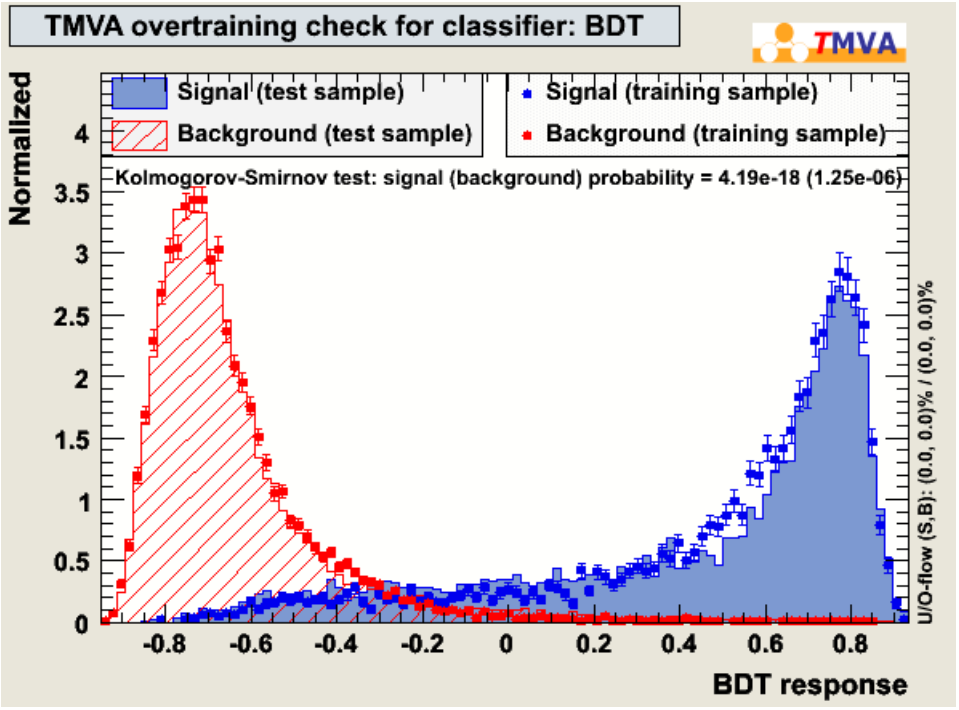

FIG. 1: BDT distribution after training for both signal (blue) and inclusive background (red).

and grand-daughters of the B_s^0 and vertex quality of the B_s^0 and its daughters. Figures 1 and 2 show the BDT distributions for the training and test samples for the inclusive and prompt background.

The BDT cuts were determined only using the 1 fb^{-1} of RunIIa data. A narrow window around the nominal $f_0(980)$ mass was chosen to keep the signal to noise ratio high. Using a mass cut of 0.96–1.0 GeV on the $\pi^+\pi^-$ mass, the BDT cut value was chosen where both S/$\sqrt{B}$ and the signal yield were high. In this way, the BDT discriminant for both the inclusive and prompt BDT was required to be greater than 0.35.

IV. YIELD RESULTS

A clear B_s^0 peak is found when the $\pi^+\pi^-$ invariant mass is near the nominal $f_0(980)$ mass. It is expected that the B_s^0 signal can be fitted to a Gaussian distribution, which provides a fitted mean mass (μ) and width (σ) for the B_s^0 peak. Since backgrounds are large, a cut of $\pm 2\sigma$ around the fitted B_s^0 peak is used to identify the $f_0(980)$ mass peak. A clear $f_0(980)$ mass peak is observed when the $\mu^+\mu^-\pi^+\pi^-$ invariant mass is within $\pm 2\sigma$ of the fitted B_s^0 mass, see Fig. 3. To decide on a $\pi^+\pi^-$ mass window to use for this analysis, a fit to the $f_0(980)$ mass peak is performed. The $f_0(980)$ has a large width [18] and is just under the KK mass threshold. This changes the line shape from a simple Breit Wigner form, particularly for higher masses and so the $\pi^+\pi^-$ mass distribution is fitted using a functional form based on Flatté [19], convoluted with a Gaussian function, that takes into account the opening of the KK threshold. The lineshape found from fitting the $f_0(980)$ in MC is used to fit the data. A $\pi^+\pi^-$ invariant mass cut of 0.91–1.05 GeV is applied to identify $B_s^0 \to J/\psi f_0(980)$ and is shown in Fig. 4. The $B_s^0 \to J/\psi f_0(980)$ mass distribution was fit to a Gaussian signal with a background function consisting of a second-degree polynomial and a Gaussian function at lower invariant mass to take into account partially reconstructed B decays.

Using identical cuts (except for the cut on the ϕ mass), a clear $J/\psi\phi$ peak is found and is shown in Fig. 5. Since the ϕ peak is so narrow, the backgrounds are much smaller for $B_s^0 \to J/\psi\phi$.

An unbinned likelihood fit was used to determine the candidate yields in each sample. The fit to the $J/\psi f_0(980)$ mass distribution shown in Fig. 4 gives the following results (statistical uncertainties only):

$$B_s^0 \text{ mass} = 5.3747 \pm 0.0036 \text{ GeV};\ \sigma = 0.0290 \pm 0.0044 \text{ GeV};\ \ 498 \pm 74\ B_s^0 \to J/\psi f_0(980) \text{ candidates.}$$

The $\mu^+\mu^-K^+K^-$ mass distribution was fit for a $B_s^0 \to J/\psi\phi$ signal using a double Gaussian function with a second-order polynomial background. A fit to the $J/\psi\phi$ distribution shown in Fig. 5 gives the following results

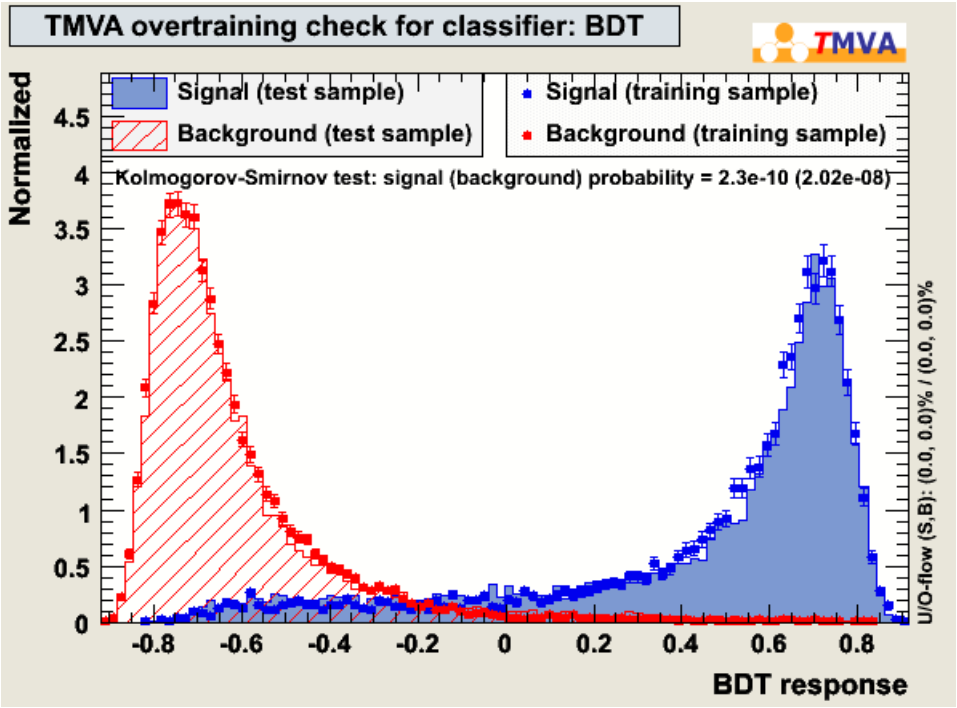

FIG. 2: BDT distribution after training for both signal (blue) and prompt background (red).

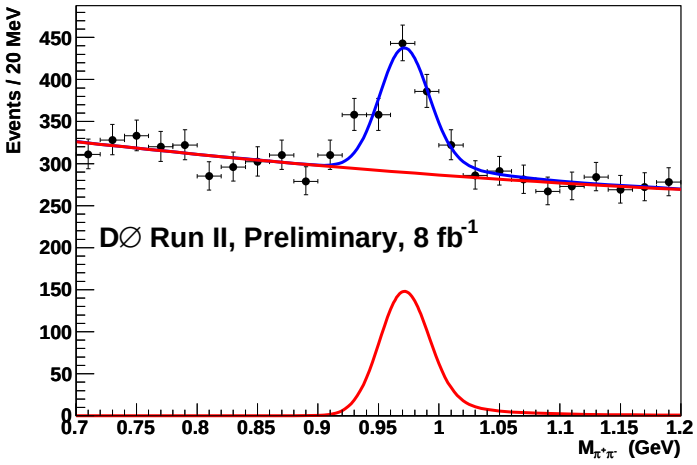

FIG. 3: $\pi^+\pi^-$ invariant mass distribution peaking at the $f_0(980)$ mass when the $J/\psi\pi^+\pi^-$ mass is $\pm 2\sigma$ around the fitted B_s^0 mass.

(statistical uncertainties only):

$$B_s^0 \text{ mass} = 5.3631 \pm 0.0008 \text{ GeV}; \;\; 2863 \pm 61 \; B_s^0 \to J/\psi\phi \text{ candidates.}$$

V. EFFICIENCIES

To determine the efficiencies of the analysis, MC signal samples were used. To take into account the effects of the instantaneous luminosity, the MC samples were overlaid with zero bias data collected during each run period. In the generation of both the $J/\psi\phi$ and the $J/\psi f_0(980)$ signal MC's, a preselection requirement of

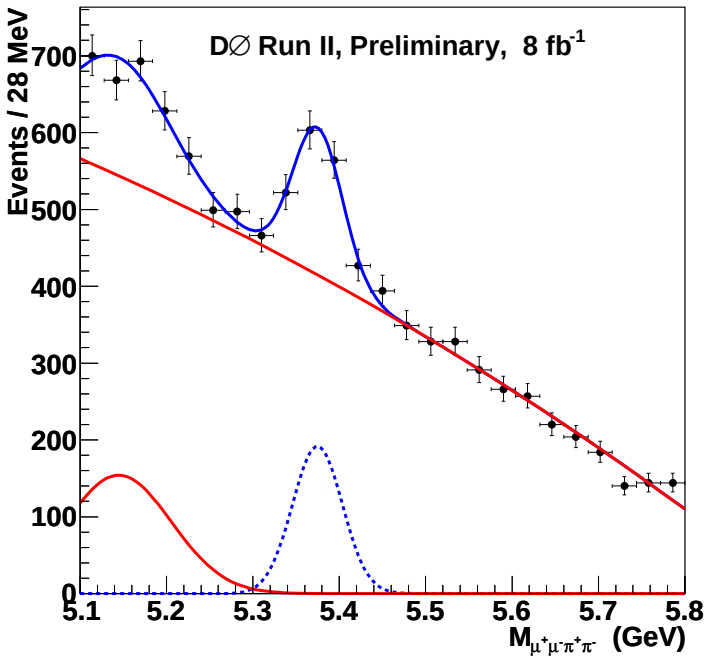

FIG. 4: $\mu^+\mu^-\pi^+\pi^-$ mass distribution peaking at the B_s^0 mass when the $\pi^+\pi^-$ mass is between 0.91 and 1.05 GeV

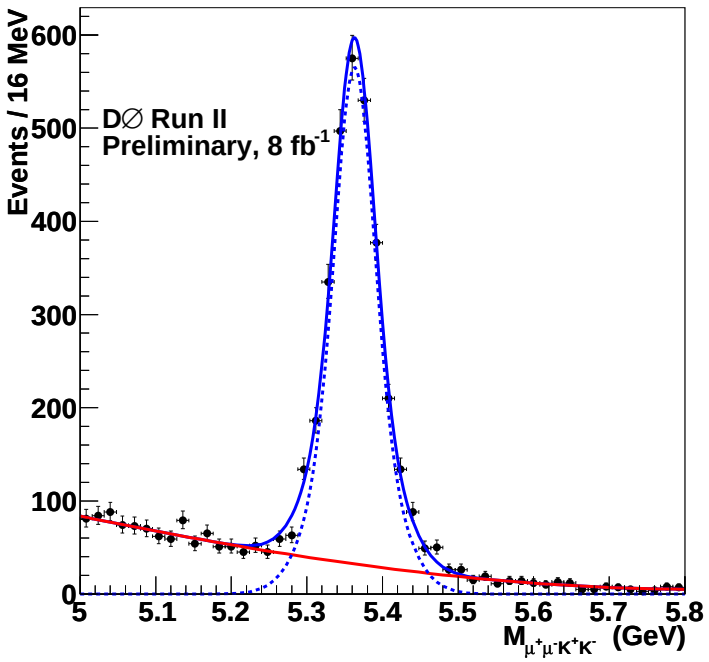

FIG. 5: $\mu^+\mu^-K^+K^-$ mass distribution peaking at the B_s^0 mass from 8 fb^{-1} of data

TABLE I: The reconstruction efficiency for $B_s^0 \to J/\psi\phi$ and $B_s^0 \to J/\psi f_0(980)$ for various running periods.

Sample	total reconstruction efficiency
$B_s^0 \to J/\psi\phi$ RunIIa	0.0231 ± 0.0004
$B_s^0 \to J/\psi\phi$ RunIIb1	0.0191 ± 0.0004
$B_s^0 \to J/\psi\phi$ RunIIb2	0.00636 ± 0.00018
$B_s^0 \to J/\psi f_0(980)$ RunIIa	0.0191 ± 0.0004
$B_s^0 \to J/\psi f_0(980)$ RunIIb1	0.0146 ± 0.0003
$B_s^0 \to J/\psi f_0(980)$ RunIIb2	0.00529 ± 0.00015

TABLE II: Reconstruction efficiencies for different run periods.

Run period	Relative reconstruction efficiency $\frac{\varepsilon_{reco}^{B_s^0 \to J/\psi\phi}}{\varepsilon_{reco}^{B_s^0 \to J/\psi f_0(980)}}$
RunIIa	1.21 ± 0.03
RunIIb1	1.31 ± 0.04
RunIIb2	1.20 ± 0.05

$p_T > 0.4$ GeV was demanded on both kaons (pions) from the $\phi(f_0(980))$. Since the p_T distributions for the pions and kaons may be different, the preselection efficiencies of this cut must be determined. To determine the preselection cut efficiencies, two additional MC sets were also generated with no p_T cuts on the pions (kaons). By comparing these two results, the preselection cut efficiencies were determined.

We found that the reconstruction efficiencies depended heavily on the MC sample used since the instantaneous luminosity was different for the various run periods, therefore we determined the reconstruction efficiencies for each run range separately. The instantaneous luminosities for runs taken during RunIIb3 were similar to the instantaneous luminosities for runs taken during RunIIb2 so the reconstruction efficiencies found from RunIIb2 were used for RunIIb3. Table I shows the results on the efficiency analysis using MC signal samples. Table I shows that the absolute reconstruction efficiencies vary in each run period, however Table II show that the relative reconstruction efficiencies are relatively stable. However, the differences in the relative reconstruction efficiency is considered a systematic uncertainty on R.

VI. SYSTEMATIC UNCERTAINTY STUDIES

A. $B_s^0 \to J/\psi\pi^+\pi^-$ background studies

One possible peaking background that affects the $B_s^0 \to J/\psi f_0(980)$ yield measurement is the non-resonant $B_s^0 \to J/\psi\pi^+\pi^-$ background. This background was studied by measuring the B_s^0 yields in $\pi^+\pi^-$ invariant mass less than the $f_0(980)$ mass. The $\pi^+\pi^-$ mass distribution from $B_s^0 \to J/\psi\pi^+\pi^-$ background where the $\pi^+\pi^-$ are non-resonant should have a much broader distribution, so determining the B_s^0 yield for lower $\pi^+\pi^-$ masses will allow a determination of the contamination in the $f_0(980)$ signal region.

In determining the $\pi^+\pi^-$ mass window to study, it is important to choose a window where one does not expect other resonances (i.e., $B_s^0 \to J/\psi K^*$). The $\pi^+\pi^-$ mass window of 0.8–0.9 GeV was chosen since in this mass range there should not be any $B_s^0 \to J/\psi K^*$ events. In fitting the distribution for any possible signal, the signal μ and σ are constrained to be the values found from the fit to the B_s^0 mass in the $f_0(980)$ signal region. The fit yields 80 ± 75 events, giving no statistically significant evidence of any $B_s^0 \to J/\psi\pi^+\pi^-$ non-resonant background, so no correction was applied.

B. Analysis cut variation

To cross check that the results do not vary with the exact value of the analysis cuts, the choice for each analysis cut was varied around its nominal value. This is an important test since the selection criteria was

TABLE III: Fractional change due to varying the exact choice of analysis cuts on the relative branching fraction

Cut	ε $(J/\psi\phi)$	ε $(J/\psi f_0)$	event yield $B_s^0 \to J/\psi\phi$	event yield $B_s^0 \to J/\psi f_0$	effect on R
BDT inc > 0.3	1.000	1.017	1.020	0.958	0.96
BDT inc > 0.4	0.993	0.980	0.975	0.945	0.98
BDT pro > 0.3	1.000	1.002	1.000	1.007	1.01
BDT pro > 0.4	1.002	1.000	1.000	0.991	0.99
$p_T(B_s^0) > 4.5$ GeV	0.997	1.000	1.000	1.000	0.99
$p_T(B_s^0) > 5.5$ GeV	1.000	0.995	1.000	0.952	0.95
$p_T(f_0(980)) > 1.0$ GeV	1.000	1.000	1.000	1.000	1.00
$p_T(f_0(980)) > 2.0$ GeV	1.000	0.987	1.000	0.980	0.99
π/K $p_T > 1.0$ GeV	1.210	1.099	1.172	1.133	1.06
π/K $p_T > 1.8$ GeV	0.724	0.771	0.797	0.744	0.88
$L/\sigma(L) > 4$	1.057	1.047	1.056	1.035	1.01
$L/\sigma(L) > 6$	0.946	0.951	0.944	0.967	1.02

TABLE IV: Effects of changing the fitting choices

Parameter	$B_s^0 \to J/\psi f_0(980)$ yield
Nominal fit (Gaussian signal + second order polynomial background with fit range 5.1–5.8 GeV)	498 ± 74
Third degree polynomial background	446 ± 72
Background function exponential+polynomial	423 ± 67
Fit range 5.1–5.6 GeV	437 ± 78
Fit range 5.15–5.8 GeV	427 ± 63
Fit range 5.05–5.8 GeV	449 ± 71

determined with 1 fb^{-1} data from RunIIa, and it is important to verify that this did not introduce a bias into the measurement. Table III shows the results from this study. As can been seen from the table, the value of R does not depend significantly on the exact choice of selection requirement.

C. Fitting cross checks

Due to large backgrounds arising from combinatorics and partially reconstructed B decays, there are significant uncertainties in the exact background shape. Therefore different parameterizations were used to describe the background and different fit regions were used to fit the data. The background polynomial was changed from a second-degree polynomial to a third-degree polynomial. The fit range was changed from the nominal 5.1–5.8 GeV and finally a different functional form for the background was used by changing the background shape to a polynomial plus an exponential.

As can be seen from Table IV, there is a fairly large variation in the number of signal events for $B_s^0 \to J/\psi f_0(980)$, indicating that the background shape is difficult to model. This fitting systematic gives the largest systematic uncertainty on R. A study was performed using same-sign pions and forming the mass distribution from $\mu^+\mu^-\pi^\pm\pi^\pm$. However, it was found the the same sign pion distribution did not describe the measured background and so could not be used to help constrain the background shape. A similar study of varying the fitting choices was performed on the $B_s^0 \to J/\psi\phi$ sample, however since the backgrounds are much smaller and easier to describe the event yield numbers changed by less than 1%.

A summary of the uncertainties on the BR are summarized in Table V.

TABLE V: Statistical and systematic uncertainties in branching fraction ratio, R

Source	Uncertainty
Statistical	0.149
Systematic from fitting	0.150
Systematic from different MC samples	0.0858

VII. FINAL BRANCHING FRACTION RATIO

The decay $B_s^0 \to J/\psi f_0(980)$ is an interesting decay mode since it can allow a measurement of the CP-violating phase in B_s^0 mixing.

A measurement of the relative branching fraction using approximately 8 fb^{-1} of data yields:

$$R = \frac{\mathcal{B}(B_s^0 \to J/\psi f_0(980); f_0(980) \to \pi^+\pi^-)}{\mathcal{B}(B_s^0 \to J/\psi\phi; \phi \to K^+K^-)} = 0.210 \pm 0.032\,(\text{stat}) \pm 0.036\,(\text{syst}).$$

The relative branching fraction of $B_s^0 \to J/\psi f_0(980)$, $f_0(980) \to \pi^+\pi^-$ to $B_s^0 \to J/\psi\phi, \phi \to K^+K^-$ should be large enough to allow a measurement of the CP-viiolating phase in B_s^0 mixin using the decay $B_s^0 \to J/\psi f_0(980)$. An analysis to measure ϕ_s using the decay $B_s^0 \to J/\psi f_0(980)$ is currently being pursued.

[1] V.M. Abazov *et al.*, (D0 Collaboration) Phys. Rev. Lett. **101**, 241801 (2008), arXiv:0802.2255 [hep-ex] and `http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/B/B61/`.
[2] D. Tonelli (CDF Collaboration), arXiv:0810.3229[hep-ex], T Aaltonen *et al.,* (CDF Collaboration), Phys Rev. Lett. **100**, 161802 (2008), arXiv:0712.2397 [hep-ex].
[3] A. Lenz and U. Nierste, JHEP 0706 (2007) 072.
[4] S. Stone, L. Zhang, arXiv:0909.5442v2 [hep-ex].
[5] S. Stone and L. Zhang, Phys. Rev. D**79**, 074024 (2009) [arXiv:0812.2832].
[6] K.M. Eckland *et.al* (CLEO Collaboration), Phys. Rev. D**80**, 052009 (2009), arXiv:0907.3201v2 [hep-ex].
[7] I. Adachi *et al.,* (Belle Collaboration) arXiv:0912.1434 [hep-ex].
[8] R. Louvot (Belle Collaboration) arXiv:1009.2605 [hep-ex].
[9] R. Aaij *et al.,* (LHCb Collaboration) Phys. Lett. **B**698:115 (2011).
[10] J. Li *et al.,* Phys. Rev. Lett. **106**, 121802 (2011).
[11] T. Aaltonen *et al.* (CDF Collaboration), arXiv:1106.3682 [hep-ex], submitted to Phys. Rev. D.
[12] A. Chandra, S. Dugad, D. Zieminska, D0 Note 4697.
[13] The proper decay length is defined as $L_{xy}(M_{B_s^0}/p_T)$, where p_T is the transverse momentum of the B_s^0, $M_{B_s^0}$ is the world average mass of the B_s^0, and L_{xy} is the transverse distance between the primary vertex and the four track vertex of the B_s^0 candidate. Primary vertices are determined by using the beamspot as a contraint and finding the vertex which contains the most tracks. Then using the other tracks not associated with the first primary vertex, but still beam constrained, search for a second primary vertex. This process is then repeated until no additional primary vertices are found. If there is more than one primary vertex in an event, the primary vertex nearest the J/Ψ candidate is selected.
[14] L. Breiman *et al.*, Classification and Regression Trees, (Wadsworth,Stamford, 1984).
[15] A. Höcker *et al.*, arXiv:physics/0703039 [physics.data-an] (2007).
[16] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).
[17] D.J. Lange, Nucl. Instrum. Methods Phys, Res. A **462**, 152 (2001).
[18] K. Nakamura *et al.*, (Particle Data Group), J. Phys. G **37**, 075021 (2010).
[19] J.B. Gay *et al.*, Phys. Lett 63B (1976) 220; S.M. Flatté, Phys. Lett. 63B (1976) 224 and Phys. Lett. 63B (1976) 228.